Onboard Science Analysis and Knowledge Discovery

Proposal for 632-07
Task Leader

Michael C. Burl, JPL, Michael.C.Burl@jpl.nasa.gov, 818-393-5345, PI

Product Description

The targeted product is a suite of algorithms that will enable both scientists and intelligent remote systems to find, analyze, and catalog spatial objects and dynamic events in large scientific datasets and real-time image streams. The algorithms will be applicable to a variety of data sources and robust to conditions that cause variations in appearance. Included will be adaptive recognizers that can be customized, without reprogramming, to look for specific object classes (e.g., volcanoes, impact craters, ice geysers, and dust storms).

The algorithms will be integrated into an innovative image data mining architecture (Diamond Eye), which will provide a low-resistance pathway for the science community to evaluate and gain confidence in this new technology. Success in the ground-based scenario will be leveraged to form partnerships with flight projects and bring advanced capabilities into the onboard setting, enabling a new era of exploration using highly autonomous spacecraft, rovers, and sensors.

This is a **continuing** task (**70% pull, 30%push**).
TRL 2-3 in FY99, TRL 3-4 in FY00, TRL 4-5 in FY01.

Benefits

The proposed product is in direct line with NASA's mission to “boldly expand frontiers in air and space”. Endowing remote platforms with the ability to recognize and respond to objects and events of interest will lead to tremendous new science opportunities that cannot be realized with existing technology. In addition to the significant impact on a broad class of future missions, development of this technology will provide scientists with tools to maximize knowledge returns from existing datasets. Specific application scenarios, as identified and described by domain scientists, appear in the accompanying letters of support.

Ground-based, PI-driven Analysis:

• Scientists will be able to directly access objects and events of interest within large datasets enabling detailed, comprehensive global analyses. Automated techniques will provide objective, reproducible feature catalogs with up to 10^3 times as many objects as would be possible with manual labeling. For crater counting, even a partial solution would offer a significant benefit, since careful human labeling of a heavily-cratered image can take ten hours or more.

• Satellite detection algorithms will provide the ability to identify subtle objects that can be easily missed by human analysts. As an example, scientists recently discovered a new satellite of Neptune that was overlooked in the Voyager imagery for ten years.

• Development of adaptive recognizers that can be easily customized for new domains will eliminate the high cost of developing many one-of-a-kind recognizers.
• Query algorithms will enable scientists to search for objects in image collections without requiring an extensive training set. Visual interfaces will allow query formation from one example or sketch.

Onboard Processing:
• Provides for a natural and powerful connection between the science PI and the spacecraft.
• By closing the loop with an onboard executive, autonomous spacecraft will be able to recognize and respond to features of interest and dynamic planetary processes such as volcanic eruptions and outgassing from comets.
• On deep-space missions characterized by long light-time delays and limited data communication bandwidth, moving the intelligence to the sensors will enable more extensive exploration and provide a dramatic improvement in the return of *science value*. Infrequent events can be captured without returning all the data. Intelligent selection and prioritization may reduce downlink by a factor of 10^6 or more with *zero* reduction in the fidelity of the returned high-resolution data.
• Specialty sensors that are too expensive to operate continuously (e.g., due to power consumption or data rate) will become practical.
• A range of algorithms will be developed to cover differing degrees of a priori knowledge
(adaptive recognizers : ad hoc queries : discovery).
• Algorithms will be reusable and applicable to a broad class of future missions.
• Outside of the basic science objective, there is synergy with other autonomy tasks such as feature detection and tracking for precision navigation and landing, vigilant monitoring for environmental hazards, and visual inspection of exposed structures, optics, and solar panels. Significant technology transfer opportunities exist with industry and other government agencies.

Technical Approach

There are clearly a number of compelling applications that would benefit from the existence of robust recognition algorithms [1]. The fundamental technical challenge is to determine what mathematical processing should be applied to the low-level, pixel representation contained in a raw image (or image sequence) in order to identify spatial objects, such as volcanoes and craters, or dynamic events such as eruptions and satellite motion. Although there is an extensive literature on object/pattern recognition, many of the techniques are not directly applicable to NASA imagery because they are targeted for man-made objects or cannot handle real-world conditions. Earlier efforts to detect natural objects in remotely-sensed imagery include the use of Hough transform [4] and matched filtering [9], but these simple techniques have not provided the necessary level of performance. There has also been significant activity in segmentation and land-use classification, but this work is not directly applicable to the problem of finding localized features and events. We outline below a comprehensive research program to address key aspects of this problem. The working model that we have adopted involves close collaboration between algorithm developers and domain scientists who help identify “hot spots”, where new computer vision, image mining, and machine learning techniques can be applied to produce the greatest science impact.

Satellite detection: Two developments this year have made the search for satellites an area of intense interest. Planetary scientists revisiting the archived Voyager dataset discovered a new satellite of Neptune that had been missed in the original analysis. Also, one of our collaborating science partners (B. Merline), using an adaptive optics technique, identified the second instance of a natural satellite orbiting an asteroid. The first such example was the Ida-Dactyl pair found in Galileo imagery. We have developed a prototype system that performs the fundamental task of identifying candidate satellites in situations in which they consist of a very small number of image pixels (perhaps only one) that barely register above background [6]. In this regime, satellites can only be detected by analysis of temporal image sequences. This is the situation as a spacecraft first approaches a known asteroid at far-field and is the most critical time to flag a new satellite so the spacecraft has time to react.

Adaptive Recognizers: Our starting point for developing adaptive recognition algorithms draws upon the extensive experience gained during development of the JARtool volcano cataloging system [3]. However, conceptual breakthroughs made since JARtool have greatly expanded the range of problems that can be addressed. One promising approach is based on the idea that the appearance of a number of real-world objects (and object classes) can be modeled as spatially-deformable configurations of lower-level parts. Recognition is achieved using local detectors to identify candidate locations for object parts and then grouping these into hypotheses which are scored based on their configuration. Spatial configurations are represented probabilistically in a special shape space in which the effects of Euclidean (translation, image plane rotation, and scaling) and affine transformations can be removed. This nonlinear approach merges the benefits of image-based techniques such as principal components analysis (used in JARtool) and feature-based techniques such as geometric invariants, alignment, and geometric hashing.

Another promising avenue involves the use of continuously-scalable and continuously-deformable template models. Starting from a prototemplate corresponding to the target object, a synthetic family of templates is generated by applying a dense set of deformations to the prototemplate. Conceptually, each family member provides a filter for detecting a specific variant of the object; however, the brute force approach of applying each filter directly cannot be used due to the computational expense. Instead, the family is compressed into a reduced set of basis functions that can be used to efficiently generate the image response to any family member. This technique extends the steerable-scalable filtering concept [5,7] and provides the foundation for our current crater detection prototype.

Ad Hoc Queries: The use of learning techniques to construct precise recognizers usually depends on an abundant supply of training examples. In many circumstances, however, it is desirable to search for objects based on a single example or even a notional concept of how an object should look, For example, the continuously-deformable template technique provides a direct method for generating queries based on a single example or sketch. The spatial configuration of parts approach, which was originally developed as a statistical model derived from training examples, can also be adapted to allow interactive query formulation. Instead of learning the probabilistic description of object geometry from real examples, the system generates cartoon-like synthetic examples and allows the user to provide feedback on the quality, which is used to update the probabilistic model. Iterating this process allows quick convergence to a usable query model.

Dynamic Events: Detection of dynamic events can benefit from both spatial and temporal analysis, although knowledge of the spatial appearance is often limited. Hence, detection of a change in appearance through analysis of an image sequence may provide the most direct and robust approach. There has been a significant amount of prior work in this area, which can be leveraged to provide a quick ramp-up in capability.

Discovery: As spacecraft venture to new environments where we don't know what to expect in advance (e.g., Pluto, subsurface of Europa, microscopic views of Mars), generic discovery algorithms that can autonomously identify ``interesting'' objects with no prior model will be required. As one example, scientists analyzing Voyager fly-by imagery of Neptune's moon Triton discovered geological features, ``ice geysers'', which had not been seen elsewhere in the solar system. It is believed within the scientific community that a detailed study of these features could lead to a totally new branch of understanding in chemistry/physics [T. Spilker, private comm.]. With an onboard discovery capability, we could envision automatically identifying and retargeting such features leading to a dramatic improvement in the scientific value of the returned data. To approach the general discovery problem, we will work from several broad assumptions: objects of interest are visually distinctive and spatially localized and repeated instances of the objects may appear in an image or set of images.

Leverage

The algorithms developed within the CETDP program will be integrated into an innovative image data mining architecture, Diamond Eye [2,8], which will provide a low-resistance pathway for the science community to evaluate and gain confidence in this new technology. The Diamond Eye concept enables scientists to interact with one or more data mining servers through a cross-platform Java applet interface. Each data mining server interprets client requests and interfaces with image sources and system resources. A computational engine associated with the server provides parallel execution of high-cost image mining algorithms. Knowledge extracted from the images is retained in an object-oriented database, which can be queried by users. The development of the Diamond Eye software infrastructure is being funded through the Office of Space Science, Applied Information Systems Research Program [J. Bredekamp], and will enable algorithms developed through CETDP to benefit through increased exposure and availability to the science community.

We are also participating in discussions with Prof. Joel Burdick and Alan Bond of Caltech regarding a seed project that will couple high-level visual recognition capabilities with a mobile agent (robot) that can actively explore its environment. This project will bring ground-based capabilities closer to the onboard setting where real-time analysis of image streams is fundamental. A target demo involves the use of the Diamond Eye interface to enable a remote user to interactively form a query that can be uploaded to the agent. Effectively, the remote user says, “Look for one of these and tell me when you find one.” We believe this will closely mirror the way in which a PI might interact with future highly-intelligent, autonomous exploration platforms.

Status

A novel algorithm for detecting objects at a continuum of scales has been developed and applied to the problem of automated crater detection. Prototype versions of the software were delivered to B. Merline and W. Colwell, planetary scientists at Southwest Research Institute, who are conducting the first systematic evaluations of the algorithm. Some experiments have shown the ability of the system to generalize across missions and planetary surfaces. Specifically, a crater recognizer derived from Viking imagery of Mars was successfully applied to a region of the moon imaged at 4-m resolution by Ranger. However, tests on more complicated terrain such as the heavily fractured Europa surface, have not been as successful. Based on insights gained in these experiments, we will continue to refine the algorithms and then proceed to another round of scientist evaluation. By tightening the interaction with domain scientists, we expect to make rapid progress toward a robust solution.

The satellite detection software was upgraded in preparation for a soft real-time analysis of the NEAR data collected at Eros. However, due to a mission glitch, the Eros dataset was not as complete as expected and no satellites were detected. In February of next year, NEAR will return to Eros enabling a more thorough search to be conducted. There is also a proposal in the works to adapt the satellite detection code to conduct a search for Vulcanoids from the Space Shuttle. (Vulcanoids are small objects believed by some to be orbiting the Sun inside of Mercury's orbit.)

Milestones

[FY 2000:] Evaluation of satellite detector code in soft real-time tests that simulate the onboard science scenario by processing downlinked NEAR data from the Eros encounter. Refinement of crater detection algorithms to provide robust intelligent-assistant capabilities on benign terrain. Continued evaluation of algorithms over a variety of datasets. Demonstration of query-by-sketch capability. Proof-of-concept demonstration of a generic focus of attention algorithm, which will serve as a seed for a discovery capability.

[FY 2001:] Satellite detector implemented as an autonomous system satisfying flight software constraints. Crater detection algorithm extended to more complex terrain and overlap conditions. System extended to perform recognition of additional spatial, as well as dynamic objects.

Customer Relevance

Future space missions will require the ability to respond quickly to important events and features as part of a general autonomous capability. Potential lander and aerobot missions such as Titan and Europa will be hampered by severe communications latency problems, which will require that spacecraft themselves be able to analyze scientific data in order to assist with site selection and data prioritization decisions. Onboard data analysis for autonomous operations is also a requirement of the Earth Science enterprise, as stressed in the recent RFI for Information Technology development in the earth sciences. Ground-based and onboard pattern recognition and data mining methods will be crucial to these goals. The HEDS strategic plan also emphasizes the role of robotic partners that can extend human exploration/analysis capabilities and continuously monitor for unexpected, potentially hazardous events.

This work has been discussed with the following customers, who have provided letters of support:

- Clark Chapman, Institute Scientist, Southwest Research Institute
- Alan Stern, Director, Department of Space Studies, Southwest Research Institute.

Biography:

Michael C. Burl is a Senior Staff Member/Technical Group Leader in the Machine Learning Systems Group at the Jet Propulsion Laboratory. He received his Ph.D. in electrical engineering from the California Institute of Technology with a dissertation entitled ``Recognition of Visual Object Classes". His research interests are focused on solving core theoretical and applied problems in computer vision, machine learning, and related fields with an emphasis on developing large-scale systems for real applications. Specific areas of interest include algorithms for finding objects in images, intelligent tools for content-based retrieval, and smart sensors/systems for autonomous exploration and data gathering. In 1996 he received a NASA Group Achievement Award on behalf of the JARtool development team for ``outstanding achievements in design, implementation, and deployment of an advanced analysis system for data mining large image collections." He previously worked in the Battlefield Surveillance Group at MIT Lincoln Laboratory where he developed algorithms for the detection and classification of tactical and strategic ground vehicles in high resolution, polarimetric SAR imagery. He has been a program committee member for various data mining and knowledge discovery conferences and has served as a reviewer for PAMI, CVPR, ICCV, ECCV, KDD, DMKD, and AES.

Selected Technical References:

[1] M.C. Burl, C. Fowlkes, J. Roden, “Mining for Image Content”, Invited paper, Joint Conf. on Systemics, Cybernetics, and Informatics / Information Systems: Analysis and Synthesis, Orlando, FL, (Jul 1999).
[2] M.C. Burl, C. Fowlkes, J. Roden, A. Stechert, and S. Mukhtar, “Diamond Eye: A Distributed Architecture for Image Data Mining”, In SPIE AeroSense - Data Mining and Knowledge Discovery, Orlando, FL, (Apr 1999).
[3] M.C. Burl, L. Asker, P. Smyth, U. Fayyad, P. Perona, J. Aubele, and L. Crumpler, “Learning to Recognize Volcanoes on Venus”, Machine Learning Special Issue on Knowledge Discovery and Applications, Vol. 30, No 2/3, pp. 165-194, (Feb/Mar 1998).
[4] A.M. Cross, “Detection of Circular Geological Features using the Hough Transform”, Int. J. of Remote Sensing, 9(9):1519-1528, (1988).
[5] W. Freeman and E. Adelson, “The Design and Use of Steerable Filters”, PAMI, v.13, 891-906, (1991).

[6] V. Gor, P. Stolorz, W. Colwell, and B. Merline, “Distant Autonomous Recognizer of Events (DARE)”, To appear in AIAA Space Technology Conference and Exposition, (Sep 1999).
[7] P. Perona, “Deformable Kernels for Early Vision”, PAMI, 17(5), 488-499, (1995).
[8] J. Roden, M.C. Burl, and C. Fowlkes, “The Diamond Eye Image Mining System”, Demo to appear at Scientific and Statistical Database Management Conference, Cleveland, OH (Jul 1999).
[9] C.R. Wiles and M.R.B. Forshaw, “Recognition of Volcanoes using Correlation Methods”, Image and Vision Computing, 11(4):188-196, (1993).

22 June 1999

Dr. Michael C. Burl
MS 126-347
Jet Propulsion Laboratory
Pasadena, CA 91109

Dear Mike:

I want to write emphasizing both the great need for, and the high payoff potential of, the kind of science data analysis autonomy work you have been doing and are proposing to continue.

At present all planetary mission spacecraft carry instruments that, though often sophisticated in terms of their data collection capabilities, are “dumber than a stump.” That is, the instruments are unaware of the science content of the data they collect. As a result, the scientific payoff from planetary missions is reduced significantly from what we could achieve with even modest degrees of onboard, intelligent data perusal. This is a particularly important loss for flyby missions, where we have no “second chance” to re-examine the target body once investigators have seen the data. It is in a real sense both an embarassing and a hugely inefficient waste of resources to send spacecraft across the solar system at great expense, only to have them take data according to often incorrect or naive pre-encounter assumptions.

Your work is opening up the possibility of remedies to this unfortunate situation. In particular, with demonstrations ranging from real-time satellite detection to surface unit analysis, to spectral analysis, a broad suite of demonstration tools are being developed at increasing TRLs, with what I and many others hope will yield near-term demonstration on upcoming technology missions. Such demonstrations would then open the door to actual application on Discovery missions to comets and asteroids, Mercury flybys, the Pluto-Kuiper Express mission, and perhaps even the Europa Orbiter.

I am personally excited by the work being undertaken to examine spectra on-the-fly for their scientific content, as this will enormously enhance our capabilities for cometary science on missions I am involved in as a PI, like Rosetta. (Comets are highly time-dynamic; as a result we must be able to follow up on discoveries in spectra before conditions change -- something that delays inherent in data transmission to the ground usually obviate).

I am also very excited about the Vulcanoids detection work, for beyond its own potential for scientific discovery, it opens great possibilities for detecting small asteroids in the Kuiper Belt objects from spacecraft that might someday actually be able to retarget themselves.

I hope very much that the work your group is involved in can continue as vigorously as it has in the past. To my knowledge, your group is pulling us faster toward the bright future of intelligent spacecraft than any other. And that is a future which I and my scientific colleagues in planetary science sadly otherwise lack, and therefore eagerly anticipate.

Sincerely,

Alan Stern
Director, Dept. of Space Studies
Southwest Research Institute

22 June 1999

Dr. Mike Burl
Information and Computing Technologies Research
Jet Propulsion Laboratory
4800 Oak Grove Dr., M/S 126-347
Pasadena, CA 91109

Dear Mike:

I would like to endorse your continued efforts to develop new algorithms that permit automatic detection of craters in imaging of surfaces of planets, satellites, asteroids, and other solar ssytem bodies. Craters are a nearly ubiquitous landform on such bodies. Since they form by fairly well understood hypervelocity impact processes, their initial shapes and topographies ("fresh" morphologies) are a given, so their subsequent evolution reflects the variety of processes that shape landforms in general. Thus they serve to calibrate studies of most planetary topography. Spatial densities of craters also form the primary basis for assessing the relative and absolute ages of geological units on planetary surfaces.

Given the enormous numbers of craters that have been imaged in the past, and will be imaged in the future by forthcoming spacecraft missions, it would be invaluable to have semi-automated or automated approaches to studying the statistics of crater populations in image archives. As you know, only a small fraction of craters in archives have been cataloged or analyzed. As a member of the imaging teams of the Galileo and NEAR missions, I have seen a wealth of scientifically important images of crater fields being returned, yet manual tabulation is slow and tedious. NEAR's prime mission at Eros, beginning next February, will provide an especially rich dataset on cratering of this never-before-explored body, which could be greatly enhanced by application of your algorithms.

In addition, one can envision the algorithms becoming sufficient robust that they could be employed automatically on-booard spacecraft for remotely analyzing images in order to make choices about measurement sequences to be employed. For instance, during a flyby without closed loop with the ground, a spacecraft could automatically assess cratered terrains and select heavily cratered (old) areas as well as sparsely cratered (young) areas for analysis.

The problem seems to be a tractable one and good progress has been made so far. I hope that it will continue.

Yours truly,

Clark R. Chapman
Institute Scientist